

**287 039**

63-1-35

# Office of Naval Research

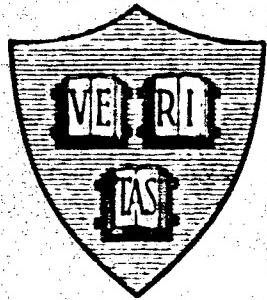
# Contract Nonr-1866(16)

NR-372-012

55282035

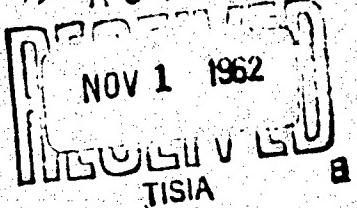
ASTRON ELE COPY

# COMPARISON OF DIFFERENT MODES OF REPRESENTATION OF MULTIVARIABLE SYSTEMS



**By**

Kumpati S. Narendra and Lyle E. McBride, Jr. ASTIA



March 30, 1962

## Technical Report No. 356

Reproduced From  
Best Available Copy

## **Crust Laboratory**

# Harvard University

# Cambridge, Massachusetts

#if. 60

Cambridge, Massachusetts  
19990610145

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

**Office of Naval Research**

**Contract Nonr-1866(16)**

**NR-372-012**

**Technical Report**

**on**

**COMPARISON OF DIFFERENT MODES**

**OF REPRESENTATION OF MULTIVARIABLE SYSTEMS**

**by**

**Kumpati S. Narendra and Lyle E. McBride, Jr.**

**Technical Report No. 356**

**March 30, 1962**

The research reported in this document was made possible through support extended Crust Laboratory, Harvard University, jointly by the Navy Department (Office of Naval Research), the Signal Corps of the U. S. Army, and the U. S. Air Force, under ONR Contract Nonr-1866(16). Reproduction in whole or in part is permitted for any purpose of the United States Government.

**Crust Laboratory**

**Harvard University**

**Cambridge, Massachusetts**

Comparison of Different Modes of Representation of Multivariable Systems

by

Kumpati S. Narendra and Lyle E. McBride, Jr.

Gordon McKay Laboratory, Harvard University

Cambridge, Massachusetts

ABSTRACT

A multivariable system may be conveniently represented either by a matrix of transfer functions or by a set of first-order differential equations. The mode of representation depends on the nature of the problem and the constraints involved. The paper deals with the relative merits of several methods as well as their inherent limitations. Concepts such as "controllability," "observability," "structure" and "interaction" which are peculiar to multivariable systems are examined to form a basis for the comparison.

y

I. Introduction

The extension of single-variable control theory to linear systems having several inputs and outputs by defining suitable input and output vectors and transfer matrices has been considered by many authors. Such matrices, defining the over-all characteristics of the system, are logical extensions of familiar frequency domain methods used in the case of single-variable systems.

The inadequacy of the transfer matrix approach to define completely even the terminal behavior under certain conditions was pointed out by Freeman [1]

and Mesarovic [2]. The latter used the concept of structure to develop different subsystem interconnections to achieve the same terminal behavior.

More recently, in the study of optimal control, interest has shifted to the time domain, and the state vector method involving a set of first-order differential equations has been used for the representation of the multivariable system. Terms such as "controllability" and "observability" have been defined in this context to describe pertinent characteristics.

The various concepts such as "structure," "interaction," "controllability," "observability," etc., are peculiar to multivariable systems and the aim of the paper is to indicate the relation between these concepts as well as the advantages and limitations of the two methods. At this time, when the synthesis of multivariable systems is still in its infancy, it is felt that a discussion of the various methods of representation and the problems to which each is best suited will facilitate further work in the field.

## II. Methods of Representation

### a. The Transfer Matrix Approach to Multivariable Systems

Over the last quarter-century a considerable body of knowledge about linear control systems has been built up. Frequency response, transient response and stability criteria, pole-zero, root-locus and compensation methods, along with many other techniques, have made possible the analytic design of even extremely complex controls for dynamic systems.

These tools are based on the concept of a transfer function which uniquely describes a real physical system under certain conditions of operation. Among the conditions which must be fulfilled if the transfer function representation is to be accurate are the following:

(1) The system must be approximately linear around its operating point so that it can be described by a linear, stationary differential equation transformable into an algebraic equation in the Laplace transform variable  $s$ .

(2) The system must be operated in such a manner that a signal enters the system or a subsystem at a prescribed input terminal, and leaves only at the output; information passes through the system in only one direction, and a single transfer function describes a single-input, single-output system.

(3) Only problems in which all initial conditions are zero or of no importance (as in steady-state behavior) can be treated by the transfer function method; in deriving the transfer function all information about the effect of initial conditions which is inherent in the differential equation is discarded. (This does not imply that the Laplace transform equivalent of a differential equation contains less information than the equation itself; it is only in obtaining the transfer function that the terms representing the effect of initial conditions are dropped.)

Transfer function techniques have been extended to systems having multiple inputs and outputs by defining suitable input and output vectors and transfer matrices. Such matrices have been used in determining the stability of multi-variable systems, in designing controllers to achieve desired dynamic performance and in specifying the conditions under which a desired transfer characteristic is realizable.

The basis of this extension is to define a vector  $X(s)$  consisting of the Laplace transforms of all the input variables in a given order, a vector  $Y(s)$  consisting of the transforms of all the output variables similarly ordered, and a transfer matrix  $A(s)$  such that the matrix equation

$$Y = AX \quad (1)$$

$$\text{where } Y = \begin{bmatrix} y_1(s) \\ y_2(s) \\ \vdots \\ y_n(s) \end{bmatrix}, \quad X = \begin{bmatrix} x_1(s) \\ \vdots \\ x_m(s) \end{bmatrix}, \quad A = \begin{bmatrix} a_{11}(s) & \dots & a_{1m}(s) \\ \vdots & \ddots & \vdots \\ a_{n1}(s) & \dots & a_{nm}(s) \end{bmatrix}$$

is valid for arbitrary input function combinations.

A system represented by a transfer matrix  $A$  may be considered to be a combination of single-variable systems interconnected in the manner of Fig. 1.

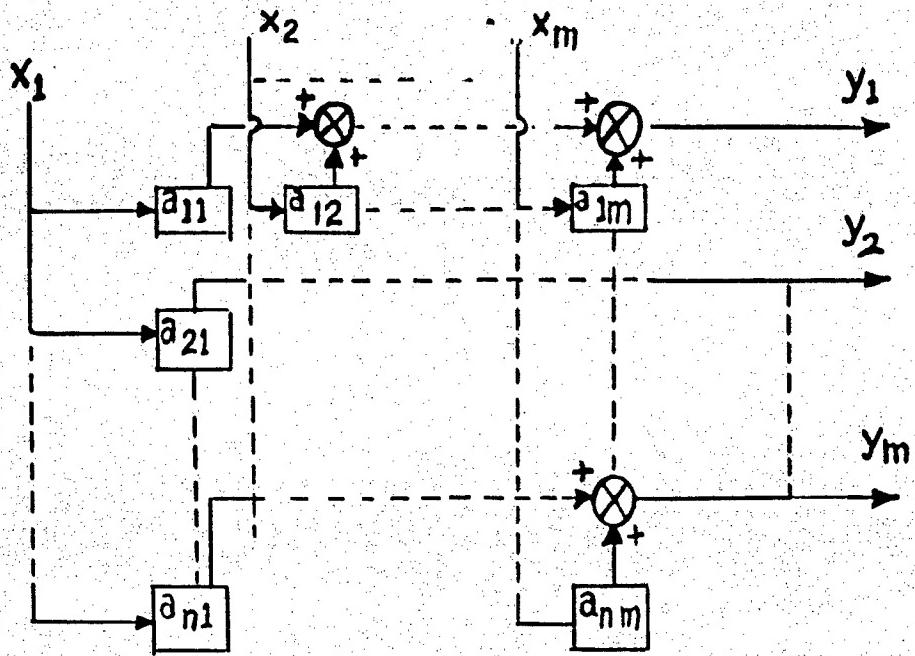


Figure 1

This combination of single-variable subsystems is now a multivariable system which obeys Eq. 1. If any combination of input functions  $x_i(s)$  is given, the resulting outputs  $y_i(s)$  can be found from this equation, since  $a_{ij}$ 's are fixed

characteristics of the system. If the  $x_i$ 's are the only variables directly affected by agencies outside the system, and the  $y_i$ 's are the only variables observed, the matrix A contains the same information about the multivariable system as the transfer function does for a single-variable system.

However, the multivariable system is found to possess properties of interest which do not exist in a single-variable system. Some of these have been discussed in the literature under the names of interactions, interrelations and Intercoupling.

(1) Input interaction. This quality, which may be roughly defined as the extent to which all inputs affect all outputs, is a function of the conventional transfer matrix A only. Complete input non-interaction is defined by Boksenbom and Hood [3] as the conversion of the n-input, n-output multivariable system into n single-variable systems; in other words, in a non-interacting system each input affects one and only one output. Non-interaction is demonstrated in the matrix representation by a diagonal A matrix (or more generally, one in which each row and column contains only one non-zero component).

(2) Output interaction. This property of multivariable systems has been discussed from various points of view, usually with a desire to eliminate it as far as possible. Output non-interaction is referred to by Freeman [1] as "independent output restoration," and is described in terms of the reaction of the system to a non-zero initial condition at output  $y_i$ , when all inputs are identically zero and all other outputs are initially zero. If under such conditions the disturbance following the "turning on of power" in the system is confined to output  $y_i$ , the system is considered non-interacting with respect to that output.

Mesarovic [2] defines a similar property called "interrelation" or "interaction," as the reaction at other outputs to an external disturbance applied to output variable  $y_i$ .

Because of conditions (2) and (3) which limit the transfer function to problems involving unidirectional signal flow in the absence of initial conditions, Eq. 1 does not contain any information about the reaction of the system to either a disturbance at the output or a non-zero initial condition. To specify output interaction, therefore, additional information not contained in the transfer matrix is necessary.

(3) Output dependence. This term is intended to describe the characteristic called "Interrelations based on External Changes of the System" in reference [2]. It describes the possibility of obtaining any independent set of desired output functions by suitable manipulation of the input functions. It may therefore be said that the outputs are independent if matrix A is non-singular when  $m = n$ . If  $m < n$  it is apparent that at least  $n - m$  outputs must be dependent on the others. In fact, in general, if the rank of A (the order of the highest non-zero determinant in A) is equal to n, the n outputs can be independently varied; if the rank of  $A + q = n$ , then q outputs can be expressed as functions of the other  $n - q$ . We could then define the case where  $q = 0$  as output independence, and all other cases as output dependence of the qth-order. (Here output dependence is equivalent to "infinite intercoupling" in reference [2].)

The extreme case of output independence is clearly the same as the case of input non-interaction; the outputs are most independent (if the superlative is permissible here) when each depends only on a single input.

#### b. "Structure" of a Multivariable System

The inability of the transfer matrix A to resolve the question of output interaction led the authors of references [1] and [2] to introduce the concept of structure.

Although their points of view differ slightly, both Freeman [1] and Mesarovic [2] define "canonical structures" which are specific ways of sub-

dividing a multivariable system into single-input, single-output subsystems described by transfer functions. Figure 1, for example, represents the "P-canonical structure" of reference [2]; various other combinations, called V-canonical, H-canonical and non-canonical structures, are shown by Mesarovic to obey Eq. 1 equally well, though the transfer functions of the component single-variable systems are no longer simply the components of the matrix A.

A proper interpretation of the structure is made to yield some information about the presence or absence of output interaction; if it is assumed that all of the blocks of Fig. 1 are ideal single-variable systems which are capable of transmitting information in only one direction, then a glance at the diagram indicates that no disturbance at  $y_1$  can be transmitted to any other output. The P-canonical structure is then said to be output non-interacting.

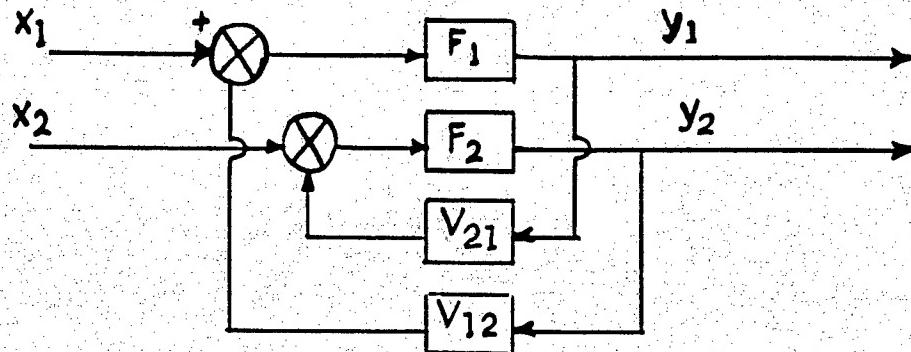


Figure 2

The V-canonical structure, on the other hand (e.g., Fig. 2) is one in which every output variable interacts with every other. A disturbance (or initial condition) at  $y_1$  will obviously be transmitted to  $y_2$  by way of the transfer functions of  $V_{21}$  and  $F_2$ . Such a structure is described as "completely interacting."

A similar use of structure, or the subdivision of multivariable systems into single-variable blocks, has arisen in connection with the concept (see Section II.c) of controllability. For example, Kalman [4] discusses the uncontrollable system shown in Fig. 3.

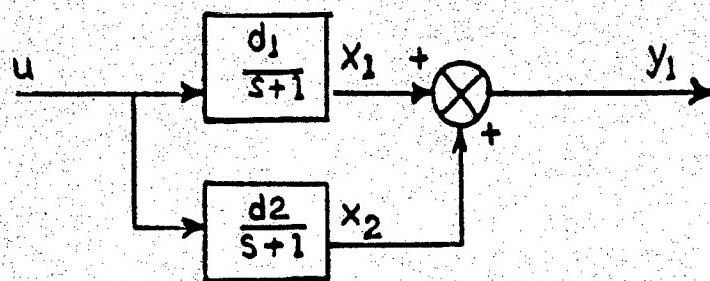


Figure 3.

If this diagram is interpreted as representing the state variable equations

$$\dot{x}_1 = -x_1 + d_1 u \quad (2)$$

$$\dot{x}_2 = -x_2 + d_2 u$$

it is apparent that  $d_2 \dot{x}_1 - d_1 \dot{x}_2 = -(d_2 x_1 - d_1 x_2)$  and that the combination  $d_2 x_1 - d_1 x_2$  represents an uncontrollable state variable (one which cannot be altered regardless of the control applied). On the other hand, since the system of Fig. 4 which is identical in transfer function has only one state variable and is completely controllable, an apparent association between structure and controllability has been established. Gilbert [5] gives additional examples of structures which are not completely controllable; he also emphasizes the fact that the transfer matrix does not necessarily indicate the correct order of the system by drawing structures of different orders which have the same transfer matrix.

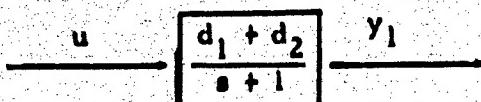


Figure 4

### c. The State Vector Method

The frequency domain approach using the concept of transfer function, described in the preceding sections, has been used in the design of control systems for over two decades. In spite of its wide application the method suffers from several inherent limitations. The tenuous relationship existing between frequency and transient responses of a system results in the designer having only qualitative information about transient response. Hence, when the transient response is rigidly specified the frequency domain method is not usually suitable. The limitations of the method also become obvious when dealing with non-linear or time-varying systems. When time responses are of interest, it is found that a direct investigation of the behavior of the differential equations governing the system directly yields more pertinent results. This has resulted in the "state vector approach" wherein the differential equations describing the physical system are studied and the system behavior is controlled directly in the time domain.

In the state vector method the multivariable system is represented by a set of differential equations of the form

$$\dot{x}_i = f_i(x_1, \dots, x_n, u_1, u_2, \dots, u_r, p_1, \dots, p_s, t) \quad i = 1, \dots, n \quad (3)$$

$$x_i(0) = C_i$$

or in vector notation

$$\dot{X} = F(X; U; P; t) \quad X(0) = C.$$

The  $x_j$  represent "n" state variables which completely specify the state of the plant at any instant,  $U_K(t)$  are the reference inputs to the system and  $P_i(t)$  the unintentional disturbances entering the system. The coordinate system with  $x_1, x_2 \dots x_n$  as coordinates spans an n-space which is called the state space. The vector  $X$  whose components are  $x_j$ , the state variables, is called the state vector and the curve traced by the state vector with the passage of time is known as a trajectory. It is worthwhile remembering that the mathematical model (represented by Eq. 3) is only an assumption. The simplification introduced by the assumption is that the behavior of the system for  $t \geq t_0$  is completely determined once ~~the initial conditions  $X(t_0)$  and  $U$  and  $P$~~  are specified. In general, the problem of control using this approach may be considered as the problem of determining the inputs  $U$  to the system such that the output  $X(t)$  corresponds as closely as possible to some desired behavior.

If we confine our attention to the problem of control of systems which are linear, the differential equation of the system may be written as

$$\dot{X}(t) = A(t) X(t) + B(t) U(t) + F(t) \quad (4)$$

where  $X$  and  $F$  are n dimensional vector functions,  $A$  is an  $(n \times n)$  matrix function and  $B$  an  $(n \times r)$  matrix function. The very first question that arises in the analysis of such a system is whether or not the state of the system can be controlled arbitrarily by the application of a suitable control function  $U(t)$ , i.e., whether any state of the system can be driven to any other desired state, by suitable control action. Such a system has been defined to be controllable [4]. It is obvious that in general controllability is a prerequisite to any type of optimal control.

The solution of Eq. 4 represents the behavior of the state of the system, and is defined by a set of functions  $x_j(t)$ . Thus  $X(t)$  implies

$$X(t; X_0, t_0) \quad (5)$$

where

$$X(t_0; X_0, t_0) = X_0$$

and  $x_j(t)$  are the components of  $X(t)$  in the state space. If the functions  $A(t)$ ,  $B(t)$  and  $U(t)$  are defined for all  $-\infty < t < \infty$  and are bounded for each  $t$ , the solution

$$X(t) = \Phi(t, t_0) X_0 + \int_{t_0}^t \Phi(t, \tau) B(\tau) U(\tau) d\tau + \int_{t_0}^t \Phi(t, \tau) F(\tau) d\tau \quad (6)$$

where  $\Phi(t, t_0)$  is the principal matrix solution of

$$\dot{X} = A(t) X(t)$$

and is called the transition matrix of Eq. 3. In the absence of external disturbances  $F(t)$ , Eq. 4 reduces to

$$X(t) = \Phi(t, t_0) X_0 + \int_{t_0}^t \Phi(t, \tau) B(\tau) U(\tau) d\tau \quad (7)$$

and when  $u(t)$  is a scalar to

$$X(t) = \Phi(t, t_0) X_0 + \int_{t_0}^t \Phi(t, \tau) B(\tau) u(\tau) d\tau \quad (8)$$

where  $B(t)$  is an  $n$  dimensional vector function. A state  $(X_0)$  at time  $t_0$  may be considered to be controllable if there exists a control  $U(t)$  which transfers the system to the state  $X = 0$  at some time  $t_1$

$$X(t_1) = 0 = \Phi(t_1, t_0) X_0 + \int_{t_0}^{t_1} \Phi(t_1, \tau) B(\tau) U(\tau) d\tau \quad (9)$$

so that  $X_0 = \int_{t_0}^{t_1} \Phi(t_0, \tau) B(\tau) U(\tau) d\tau$  holds for some  $U(t)$ . The mathematical

theory of the controllability of linear dynamical systems is discussed in considerable detail in [6] and the conditions for controllability of time-varying and constant systems are derived.

Considering the case when  $A$  and  $B$  in Eq. 4 are constant, the system is completely controllable if, and only if, the constant  $n \times rn$  matrix  $\{B, AB, A^2B, \dots, A^{n-1}B\}$  has a rank "n." This ensures that the control function  $U(t)$  operated on by  $B$  spans the entire state space, so that each state variable may be independently affected by the control action. (For a mathematical derivation of the necessary and sufficient conditions of controllability as well as other mathematical concepts the reader is referred to reference [6].)

Considering the time-invariant case the equation describing the system is of the form

$$\dot{X} = AX + BU \quad (10)$$

when  $U$  is a vector, and

$$\dot{X} = AX + Bu \quad (11)$$

when  $U$  is a scalar quantity. Since by definition of controllability it is clear that it is invariant to any non-singular linear transformation, if we define

$$X = PY$$

Eq. 10 becomes

$$\dot{Y} = P^{-1} APY + P^{-1} BU. \quad (12)$$

By the proper choice of  $P$  it is seen that the equation may be reduced to normal form. If in Eq. 10 it is assumed that  $X$  is already in its principal coordinates,

$A$  is in its Jordan canonical form. Assuming that the matrix  $A$  has "q" distinct eigenvalues  $\lambda_1, \dots, \lambda_q$  and an eigenvalue  $\lambda_{q+1}$  of multiplicity  $(n - q)$ ,  $X, B$  and  $A$  in Eq. 11 may be expressed in the following form

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \\ x_{q+1, 1} \\ \vdots \\ x_{q+1, n-q} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_q \\ b_{q+1, 1} \\ \vdots \\ b_{q+1, n-q} \end{bmatrix}, \quad A = \begin{bmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_2 & & & & & \\ & & & \ddots & & & & \\ & & & & \lambda_q & & & \\ & & & & & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 0 \\ & & & & & & & & \ddots & & \\ & & & & & & & & & 0 & \\ & & & & & & & & & & \ddots & \\ & & & & & & & & & & & \ddots & \\ & & & & & & & & & & & & \ddots & \\ & & & & & & & & & & & & & \ddots & \\ & & & & & & & & & & & & & & \ddots & \\ & & & & & & & & & & & & & & & 0 \end{bmatrix} \quad (13)$$

The transition matrix is of the form

$$\Phi(t) = \begin{bmatrix} e^{\lambda_1 t} & & & & & & & \\ & \ddots & & & & & & \\ & & e^{\lambda_q t} & & & & & \\ & & & \ddots & & & & \\ & & & & e^{\lambda_{q+1} t} & & & \\ & & & & & \begin{bmatrix} 1 & t & t^2 & \cdots & t^{n-q-1} \\ 0 & 1 & 0 & \cdots & 0 \\ & & \ddots & \ddots & \ddots \\ & & & 1 & 0 \end{bmatrix} & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{bmatrix} \quad (14)$$

For complete controllability,  $\Phi(t)B$  should have "n" components which are linearly independent and non-zero over each positive interval of time.

The block diagram representation of Eq. 11 is shown in Fig. 5. This in turn may be expressed in block diagram form as shown in Fig. 6.

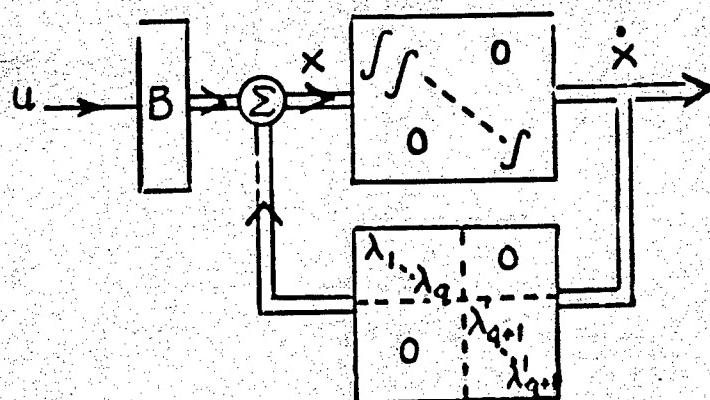


Figure 5

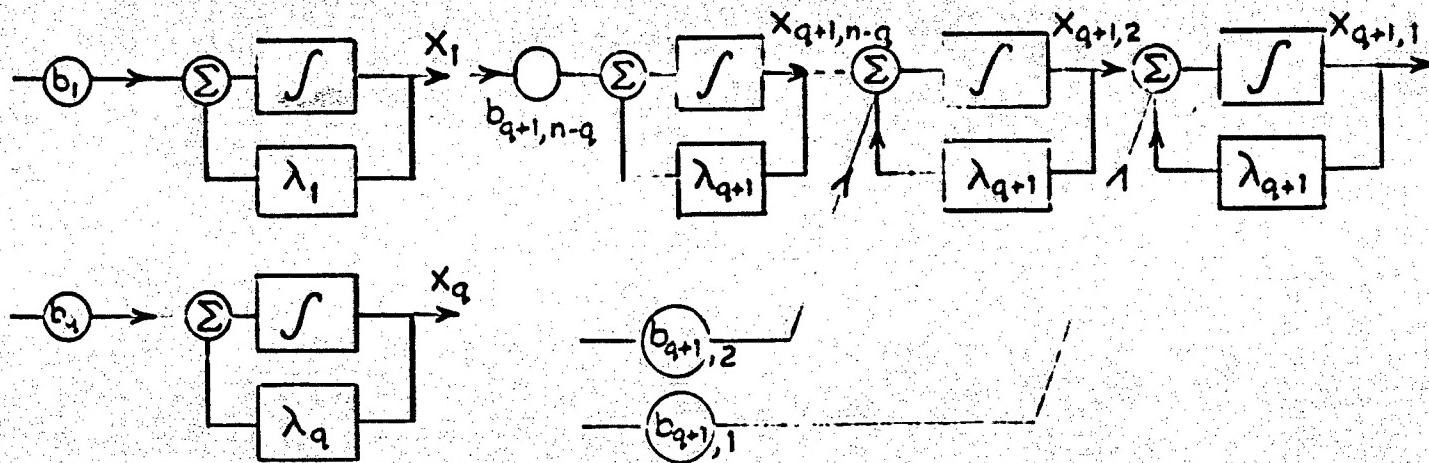


Figure 6

To make  $\Phi(t)B$  have "n" linearly independent components it is clear that  $b_1, b_2 \dots b_q$  and  $b_{q+1, n-q}$  should not be zero in Fig. 6.  $b_{q+1, n-q} \neq 0$  ensures that the state variables  $x_{q+1, 1} \dots x_{q+1, n-q}$  can be controlled independently. In other words, at least one of the control inputs should affect

$x_1, \dots, x_q$  and  $x_{q+1}, \dots, x_n$  which is expressed by the fact that the matrix  $\{B, AB, \dots, A^{n-1}B\}$  should have a rank "n."

Controllability as defined above implies that by the suitable application of a control function  $U$  the state variables may be affected independently in time so that for some value of "t" the variables  $x_i$  may be made to assume a specified value. From the above discussion it is clear that a system of the form shown in Fig. 7a is controllable while the one indicated in Fig. 7b is not.

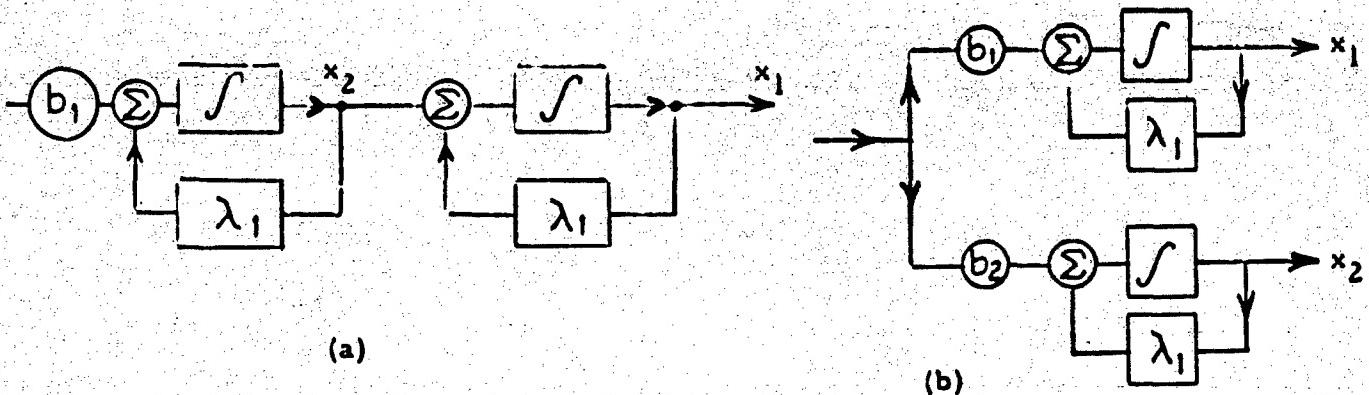


Figure 7

When the state vector representation is used for the description of a multivariable system, the problem generally becomes the determination of  $U(t)$  (subject to constraints on  $U$  and  $X$ ) to minimize or maximize a performance criterion. The state vector  $X$  is driven from an initial state  $X_0$  to a final fixed state  $X_f$  or more generally a varying state  $Z(t)$ .

A second concept [4] which is of importance while using the state vector method is termed "observability" and applies to the situation when the state

vectors are not directly accessible but have to be computed from the available data. In particular, in the linear case, if the output  $Y$  is such that its components  $y_i$  are linear combinations of the state variables, we have

$$Y = CX \quad (15)$$

where  $C$  is a  $p \times n$  matrix function and  $Y$  is a  $p \times 1$  vector. The system is "observable" if all the state variables  $x_i$  of the state vector can be determined by observing  $Y$  over a finite interval of time.

Assuming that the equations of the system are expressed in terms of the principal coordinates and

$$\dot{X} = AX + BU \quad (16)$$

and

$$CX = Y$$

and if  $A$  has  $n$  distinct eigenvalues  $\lambda_1, \lambda_2 \dots \lambda_n$  the block diagram representation of the system is as shown in Fig. 8.

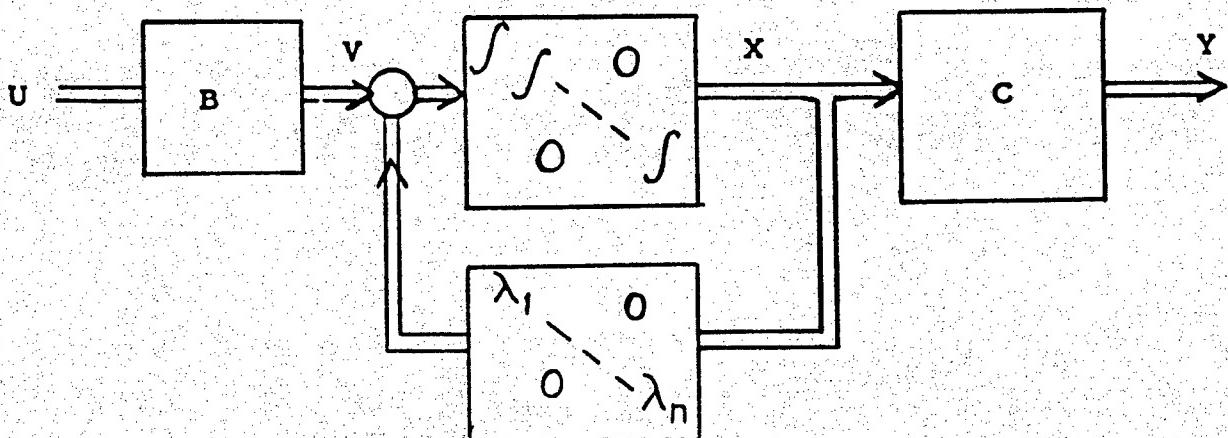


Figure 8

It is seen that each component  $v_i$  of the vector  $V$  controls the corresponding state  $x_i$ . If the  $i$ th row of  $B$  contains only zero elements,  $v_i$  is zero and  $x_i$  is uncontrollable. Similarly, if the elements of the  $j$ th column of  $C$  vanish,  $x_j$  is not observable. This definition is suggested in reference [5].

For the more complex case when the system has multiple characteristic roots such a simple definition is not valid. Due to the internal feedback in the system, Fig. 5 variations in  $x_j$  may be observed by observing the other state variables. Similarly, the system may be completely controllable even though a row of the  $B$  matrix vanishes.

### III. Relative Merits of Various Methods of Representation

#### a. The Transfer Matrix

The matrix  $A$ , which relates input and output variables in the frequency domain, is the simplest and most direct means of applying techniques developed for single-variable servomechanisms to the field of multivariable systems.

(1) Input non-interaction. In addition to the problems of controller design involving stability, steady-state errors and similar properties of the component subsystems (a number of which have been studied in the literature by means of the transfer matrix), the synthesis of a controller to produce input non-interaction from an interacting plant may be highly desirable in order to reduce the number of degrees of freedom of the designer and simplify the dynamic analysis. The requirements for non-interaction can be directly and generally obtained from the matrix  $A$ .

Let us first consider a given interacting plant of known transfer matrix  $A_p$ , which is to be connected in cascade (Fig. 9) with a forward acting controller whose transfer matrix is  $A_c$ .

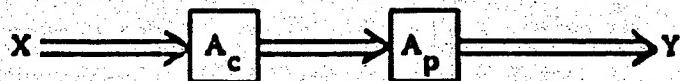


Figure 9

Then  $A_p A_c X = Y = AX$ , and  $A$ , the resulting over-all transfer function, is to be diagonal. The control elements are determined from

$$A_c = A_p^{-1} A, \quad A = [a_{ij} \delta_{ij}] \quad A_p^{-1} = [\bar{a}_{pij}] \quad A_c = [a_{cij}]$$

or

$$a_{cij} = a_{jj} \bar{a}_{pij} \quad (17)$$

hence

$$\frac{a_{cij}}{a_{ckj}} = \frac{\bar{a}_{pij}}{\bar{a}_{pkj}} \quad (18)$$

The control elements  $a_{cij}$  can be determined if  $A_p$  is non-singular.

From Eqs. 17 and 18 it is apparent that the ratios of the elements in a given column of the matrix  $A_c$  are fixed by the condition that the total system be non-interacting (that  $A$  be diagonal). However, the freedom to determine an arbitrary transfer matrix  $A$  remains, as each column may be multiplied by an arbitrary transfer function  $a_{jj}$ .

If control elements are all connected in the feedback path, the block diagram of Fig. 10 results, and

$$A_p(X + A_c Y) = Y \quad \text{or} \quad Y = (I - A_p A_c)^{-1} A_p X = AX$$

then

$$A_p = (I - A_p A_c)A \quad A - A_p = A_p A_c A$$

$$A_c = A_p^{-1} - A^{-1} \quad A^{-1} = \left[ \frac{1}{a_{ij}} \delta_{ij} \right].$$

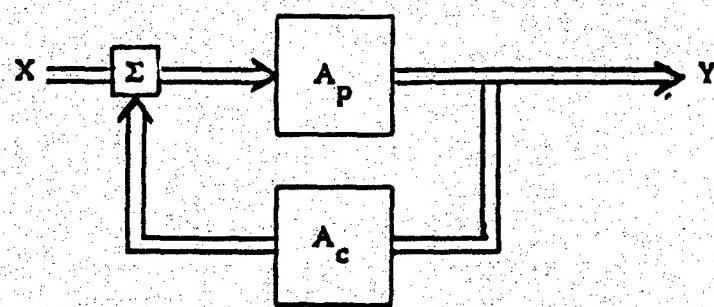


Figure 10

Then we find that

$$a_{cij} = \bar{a}_{pij} \quad \text{if } i \neq j \quad (19)$$

and

$$a_{cii} = \bar{a}_{pii} - \frac{1}{a_{ii}} \quad . \quad (20)$$

In this case, all off-diagonal elements of the controller are fixed by the non-interaction requirement, but the diagonal terms remain free to establish any desired diagonal transfer matrix  $A_c$ . Here both  $A_p$  and  $A_c$  must be non-singular.

The requirements of non-singularity are not surprising. If  $A_p$  were singular it would indicate that one or more of the outputs is not independent (is a constant function of the other outputs, regardless of the combination of inputs applied). It is clearly impossible, under these conditions, to make the outputs vary independently by changing the plant inputs. If  $A_p$  were singular, it would indicate that one output is always zero (since  $A_p$  is diagonal). This is possible with a feed-forward controller, since by a proper selection of the  $n$  inputs any number of outputs can be made zero (if  $A_p$  is non-singular). But the feedback controller would then only have  $n - 1$  non-zero inputs, and consequently

its output could contain only  $n - 1$  independent functions of the plant output, not enough to determine uniquely the system outputs.

Thus, in an  $n \times n$  system there always remain  $n$  conditions, after non-interaction has been attained which can be chosen to produce the desired diagonal transfer matrix. The use of a feedback controller has the property (which may be advantageous) that the controller components determined by the condition of non-interaction are independent of the dynamic characteristics chosen for the diagonal elements of  $A$ .

(b) The limitations of "structure." The use of structure as a necessary property in the definition of a multivariable system was first suggested in order to provide information about output interaction; if a multivariable system is really made up of ideal, isolated, unidirectional single-variable systems (as the subdivided block diagram implies), then it is clear that the system of Fig. 1 is output non-interacting while that of Fig. 2 is completely interacting.

A difficulty may arise, however, if we require quantitative information about the interaction. If all inputs are identically zero ( $X \equiv 0$ ) and a known disturbance  $y_1(s)$  is applied, what is the resulting  $y_2(s)$ ? In the system of Fig. 1, the answer is obviously  $y_2 \equiv 0$ . In the structure of Fig. 11

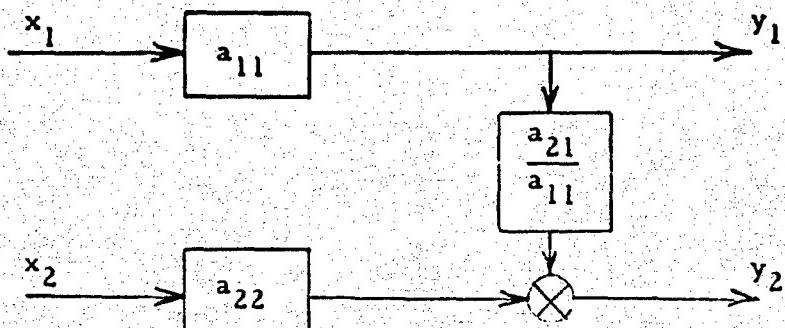


Figure 11

$y_2 = y_1 (a_{21}/a_{11})$ . But in the V-canonical structure of Fig. 2, while it seems clear that  $y_2(s) \neq 0$ , the only equations at our disposal (when  $X = 0$ ) are

$$0 = y_1 - V_{21} F_1 y_2 \quad (21)$$

$$0 = -V_{12} F_2 y_1 + y_2$$

or in matrix form  $(I - FV)Y = 0$ , which in general has only the trivial solution  
 $y_1 = y_2 = 0$ .

The problem is directly traceable to condition (2) which was stated as a requirement for transfer function representation in Section II: signal flow must be unidirectional. We have been able to violate this condition in the case of Figs. 1 and 11 only by tacitly defining the reverse transfer function of a block to be zero. As long as a given block encounters input signals at only one of its terminals, this assumption provides a solution; in Fig. 2, however, block  $F_1$  encounters an input signal simultaneously at both terminals:  $y_1$  at the output and  $y_1 V_{21} F_2 V_{12}$  at the input. Clearly, no solution is possible in the general case. In order to obtain a meaningful quantitative description of output interaction, a system description valid for bidirectional signal flow is needed. Such a representation is discussed in the following section.

A similar difficulty arises in the representation of uncontrollable (or unobservable) systems by means of a structure, or block diagram of first-order transfer functions. As mentioned in Section II, an intuitive interpretation of Fig. 3 results in the equation

$$\frac{dz}{dt} = -z \quad (22)$$

where  $z = d_2 x_1 - d_1 x_2$ .

Consider, however, the network of Fig. 12 where  $u$  is the input voltage and  $x_1$  and  $x_2$  are output voltages measured at the points indicated.

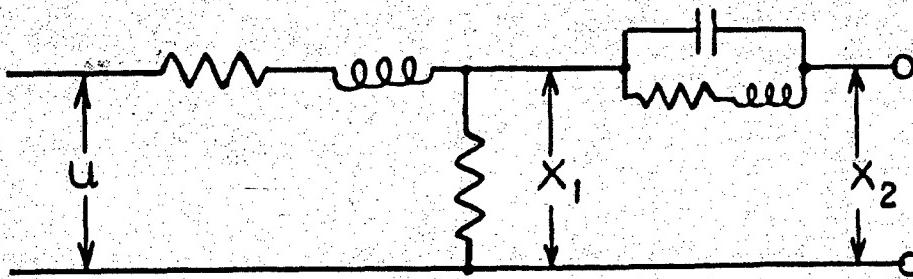


Figure 12

If the component values are properly chosen, the applicable equations are

$$x_1 + x_2 = u \quad (23)$$

$$z + k + z = 0$$

where  $z = x_1 - x_2$ . If these equations are transformed into the frequency domain, and all initial conditions are set equal to zero, the resulting transfer function equations are

$$x_1(s) = x_2(s) = \frac{1}{s+1} u(s)$$

corresponding to the block diagram of Fig. 13, which is similar to that of Fig. 3.

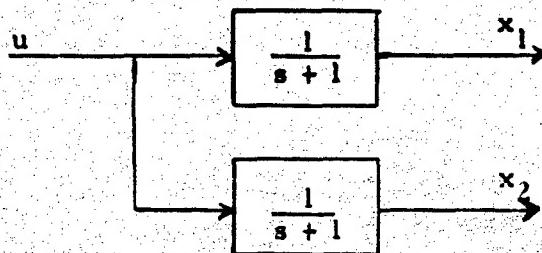


Figure 13

In this case, however, there are two initial conditions,  $z(0)$  and  $\dot{z}(0)$ , which cannot be brought to zero by any choice of  $u$ .

From this example it can be seen that, because the transfer function is valid only when all initial conditions are zero, neither a transfer function nor a block diagram made up of transfer functions can uniquely specify the number of initial conditions necessary to define a general solution to the system equations, and therefore neither the transfer matrix nor the structure can specify the order of the system, although the minimum order can be computed from the transfer matrix. The statement of reference [5] that uncontrollable or unobservable elements do not appear in the transfer matrix, applies equally to the block diagram or structure of the system.

In short, the structure or subdivision into single-variable blocks of a multivariable system gives no more information about the response of the system than does the transfer matrix. This result is well known in the case of single-variable systems.

The state vector, or differential equation representation, which describes completely the effect of initial conditions, can be used to determine whether or not output interaction exists. Since  $\phi_{ij}(t)$  is the response of  $x_i$  to a unit initial condition on  $x_j$ , complete output non-interaction in accordance with Freeman's definition corresponds to a diagonal  $\Phi$  matrix; if  $\phi_{ij}(t) \equiv 0$  an initial condition in  $x_j$  has no effect on  $x_i(t)$ .

While output interaction of this type can be expressed in terms of the  $\Phi$  matrix, if interaction is more generally interpreted as the effect of an externally imposed change in one output variable on another output the state vector representation does not provide adequate information.

### c. A Multivariable System as a Generalized Electric Network

One multivariable system which has been studied in great detail is the electric network made up of constant impedances and stationary voltage sources. In addition, by means of either the impedance analog or the mobility analog, mechanical systems can be expressed in terms of equivalent networks. Furthermore, a great many, if not all, other types of linear, stationary systems are governed by the same differential equations applicable to networks; thus a wide variety of systems, including all linear electromechanical systems, can be studied by means of network analysis.

(1) Inputs and outputs. To make a connection between the representation of the preceding section and the network representation, let us consider a system which has  $n$  inputs  $x_1 \dots x_n$  and  $m$  outputs  $y_1 \dots y_m$ .

In the above discussion it was assumed that the inputs  $x_i$  are independent variables which may be manipulated for the purpose of altering the outputs  $y_i$ ; the outputs are defined as variables which cannot be directly manipulated externally, but in whose values we are interested. It is assumed that although other variables exist inside the system they cannot be manipulated or are not of interest; therefore, they are omitted.

If the system is an electromechanical system with  $n$  terminals (or more properly  $n$  ports) it is apparent that two variables are present at each terminal. For example, a single shaft is associated simultaneously with a position and a torque; an electrical terminal similarly has a voltage and a current. Other combinations of variables may be chosen (velocity and torque, voltage and power, etc.) but the condition remains that the state of one terminal is specified by two variables. Thus an  $n$ -port network has  $2n$  variables existing at its terminals.

If, for example, the  $n$  voltages are fixed by external connections, the  $n$  currents are completely determined for a given network. In general, no current

can be altered without simultaneously changing one or more voltages. We can then say that the  $n$  voltages are independent variables, meaning that they are determined by forces outside the network; the  $n$  currents must then be dependent, determined by the network characteristics and by the  $n$  independent voltages. Of course it is not necessary that the independent variables all be voltages or all currents; any combination of voltages and currents may be considered independent, with one restriction: the voltage and current at any one terminal may not simultaneously be both independent or both dependent. This is a consequence of the general statement that voltage and current at any point are determined by the simultaneous solution of two equations, one representing the current-voltage relationship on one side, the other representing a similar relation on the other side of the point. Neither conditions inside the network nor those outside can alone determine both current and voltage.

If we now consider this network as a control system, we must classify the inputs and outputs as dependent or independent. Clearly the inputs are independent, since they are manipulated from the outside, and the outputs are dependent, since they are to be controlled by the inputs. Since, however, the inputs are not generally found at the same terminals as the outputs, the selection of  $n$  inputs and  $m$  outputs will mean that the system has  $n + m$  ports, and hence  $2(n + m)$  terminal variables. We then are left with  $m$  independent variables which are not inputs, and  $n$  dependent variables which are not outputs. These correspond to the conjugate variables at the output and input terminals. Let us call the  $m$  independent variables secondary inputs, and the  $n$  dependent variables secondary outputs.

(2) The Kron impedance mixed method. A very general method of solution of electric network problems is the Kron mixed method described in Chapter

IV of reference [7]. It is applicable in its original form to networks containing any combination of constant inductances, constant capacitances, constant resistances, and constant frequency voltage and current generators. As external perturbations or inputs, it is assumed that between any pair of terminals there may be connected either a constant voltage source or a constant current source. It is assumed that if the number of external voltage and current sources is less than the number of node pairs of the network, the remaining node pairs are connected by voltmeters. A "voltmeter mesh" may be considered equivalent to a constant current generator of zero amplitude.

If our network is to represent a system which has a transfer matrix of the type discussed in the previous section, all outputs must be zero when all inputs (primary and secondary) are zero. For the network, this condition is true only if no internal sources are present. Our basic network thus contains only impedances. This condition does not exclude mutual impedance between branches, whether symmetric, asymmetric or unilateral.

The network impedances are defined in terms of the relation  $e = z(\omega)i$ , where  $e$  and  $i$  are the complex descriptions of a sinusoidal voltage and current of frequency  $\omega$ . This, however, is commonly recognized as being a special case of  $e(s) = z(s)i(s)$  where the transform variable  $s$  is replaced by  $j\omega$ . We can, therefore, interpret the impedances as transfer functions, and after solving the network equations find the output currents, for example, for any arbitrary input form.

The notation in the following discussion is in part that of reference [7]. If there are no internal sources, the network obeys the matrix equation

$$E' = Z'j' \quad (\text{from reference [7], Eq. 44-10}) \quad (24)$$

where

$$E' = \begin{bmatrix} O_1 \\ \vdots \\ O_M \\ E_1 \\ \vdots \\ E_V \\ *V_{V+1} \\ \vdots \\ *V_P \end{bmatrix} \quad J' = \begin{bmatrix} *j_1 \\ \vdots \\ *j_M \\ *i_1 \\ \vdots \\ *i_V \\ I_{V+1} \\ \vdots \\ i_P \end{bmatrix} \quad (25)$$

$$Z' = C_t Z C$$

$Z$  is the network impedance matrix,  $C$  is the transformation matrix relating the branch currents to the mesh currents,  $C_t$  is the transpose of  $C$ ,  $E_i$  is the known voltage applied to the  $i$ th node pair,  $*V_j$  is the unknown voltage appearing across the  $j$ th node pair,  $*j_k$  is the unknown current in the  $k$ th Maxwell mesh,  $*i_l$  is the unknown current in the external voltage generator,  $E_i$  and  $I_j$  is the known current applied in the external circuit between the nodes of the  $j$ th pair. Also the subscript  $M$  is the number of Maxwell meshes in the original network,  $V$  is the number of external voltage sources applied and  $P$  is the number of node pairs of the network. ( $V \leq P$  is required by Kirchoff's laws.)

In the above equations, the starred scalar quantities are unknown, the matrices  $Z$  and  $C$ , and hence  $Z'$ , are characteristics of the network, the applied voltages and currents  $E_i$  and  $I_j$  are known and the first  $M$  elements of  $E'$  are zero because of the absence of internal voltage sources.

We can now assume that our inputs are among the applied external variables  $E_i$  and  $I_j$ , and that our outputs are among the starred unknowns. Since Eq.

24 represents P + M equations in P + M unknowns, our outputs can be solved for in terms of the inputs and other externally determined variables, but we do not have an explicit transfer matrix. Also, Eq. 24 involves a number of mesh currents which are completely internal and therefore of no interest to us in describing the transfer characteristic of the network; there may also be a number of node pairs where no external connection has been made and the voltages across these are of no interest.

In writing the vectors  $E'$  and  $J'$  above, the voltages were arbitrarily ordered so that known variables appear in a block, and the unknowns in an adjacent block; however, Eq. 24 holds regardless of which voltages are known. As long as P voltages and currents are given (any number may be identically zero) the P + M unknown voltages and currents may be found.

By partitioning the  $Z'$  matrix and the voltage and current vectors, and performing the necessary matrix operations (involving inversions which are justified in reference [7]), an equation of the following form can be obtained:

$$\begin{bmatrix} *i_1 \\ \vdots \\ *i_V \\ *v_{V+1} \\ \vdots \\ *v_P \end{bmatrix} = Q \begin{bmatrix} E_1 \\ \vdots \\ E_V \\ I_{V+1} \\ \vdots \\ I_P \end{bmatrix} \quad (26)$$

where  $Q$  is a  $P \times P$  matrix derived from  $Z'$  by partition, appropriate matrix operations on the submatrices, and reassembly.

Equation 26 represents the complete terminal solution for this particular choice of dependent and independent variables. If this choice is altered,

the manner in which  $Z'$  is partitioned changes and therefore the resultant  $Q$  is changed.

If no connection is to be made to a certain node pair (terminal) the corresponding current is set equal to zero, and the corresponding voltage is omitted as being of no interest. This eliminates one row and one column from the matrix  $Q$ . This can be continued until only the terminals in which we are interested remain. We now have a matrix equation of the form:

$$D = Q' I \quad (27)$$

where  $D$  is the vector of all dependent variables (primary and secondary outputs), and  $I$  is the vector of all independent variables (primary and secondary inputs) and  $Q'$  is the matrix  $Q$  with all rows and columns not corresponding to input or output terminals omitted.

We have now obtained a transfer function involving all terminal variables, for a linear electric network. If the equations governing a linear multivariable system are put into this form, will this give us additional information about its behavior? Note that this form can be obtained only after it has been decided which variables are independent and which are dependent. The transfer matrix, like the transfer function, remains unilateral in the sense that the role of independent and dependent variables cannot be interchanged without altering the transfer characteristic. It is, however, bilateral in the sense that a given terminal can simultaneously receive a signal and transmit one.

#### d. A Transfer Matrix Representation of a Multivariable System

Let us define the vectors. Here  $v_i$ 's are the dependent variables,  $u_i$ 's are independent variables,  $y_i$ 's are true outputs,  $x_i$ 's are true inputs,  $z_i$  is the secondary output conjugate to (existing at the same terminal as)  $x_i$ , and  $w_i$  is the secondary input conjugate to  $y_i$ .

$$V = \begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_m \\ z_1 \\ \vdots \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ \vdots \\ v_{m+n} \end{bmatrix} \quad U = \begin{bmatrix} u_1 \\ \vdots \\ \vdots \\ u_{n+m} \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \\ w_1 \\ \vdots \\ \vdots \\ w_m \end{bmatrix} \quad (28)$$

Now an equation of the same form as 27, but describing an  $n + m$  terminal multivariable system can be written as

$$V = PU \quad (29)$$

where  $P$  is the transfer matrix analogous to  $Q'$  and

$$P = \begin{bmatrix} p_{11} & \cdots & \cdots & \cdots & p_{1(n+m)} \\ \vdots & & & & \vdots \\ p_{(m+n)1} & \cdots & \cdots & \cdots & p_{(m+n)(m+n)} \end{bmatrix}$$

Now  $P$  may be partitioned into:

$$P = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \text{ where}$$

$$A = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \vdots & & \vdots \\ p_{m1} & \cdots & p_{mn} \end{bmatrix}$$

$$B = \begin{bmatrix} p_{1(n+1)} & \cdots & p_{1(n+m)} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \vdots & & \vdots \\ p_{m(n+1)} & \cdots & p_{m(n+m)} \end{bmatrix}$$

$$C = \begin{bmatrix} p_{(m+1)1} & \cdots & p_{(m+1)n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \vdots & & \vdots \\ p_{(m+n)1} & \cdots & p_{(m+n)n} \end{bmatrix}$$

$$D = \begin{bmatrix} p_{(m+1)(n+1)} & \cdots & p_{(m+1)(m+n)} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \vdots & & \vdots \\ p_{(m+n)(n+1)} & \cdots & p_{(m+n)(m+n)} \end{bmatrix}$$

Then Eq. 29 is equivalent to the two equations

$$Y = AX + BW \quad (30)$$

and

$$Z = CX + DW. \quad (31)$$

If we assume that the independent variables  $w_1$  conjugate to the outputs  $y_1$  are all equal to zero, then Eq. 30 is reduced to Eq. 1. But this is just what is ordinarily done in single-variable theory; if the output is a position, all known loads attached to the output shaft are considered to be a part of the system, and the externally applied torque at the output shaft is therefore assumed to be zero. The two independent variables are thereby reduced to one, the input. Such a system is therefore a special case of a multivariable system which has its secondary input equal to zero.

If the network described by the matrix  $P$  is the entire system under consideration, the secondary outputs  $z_1$  are of no interest, and Eq. 31 can be ignored. If, however, we wish to interconnect this network with another described by the matrix  $P'$ ,  $z_1$  acquires a new importance. Suppose the input signal  $x_1$  is to be obtained from the output  $y'_1$  of network  $P'$ . If  $x_1$  and  $y_1$  are both voltages, this result can be obtained by connecting terminal 1 to terminal 1'. This connection requires at the same time that the two currents be equal, or that  $z_1 = w'_1$ . The corresponding block diagram is:

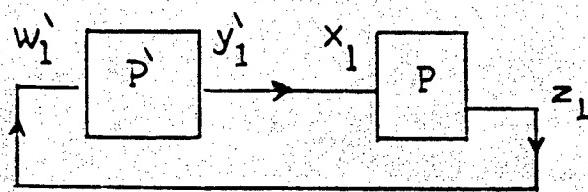


Figure 14

Now, even though only  $y_1 \dots y_n$  are outputs of interest,  $x_i$  influences  $w_i$  and hence  $y'_i$  and  $x_i$ , and Eq. 31 can no longer be ignored. This illustrates the reason why conventional transfer function blocks must be connected through isolating amplifiers or their equivalent. Isolation means that  $x_i = w_i$  is zero and that Eq. 1 is valid.

If the P matrix is written for a single-input, single-output electrical network, where it is assumed that voltage  $E_1$  is the input and the current  $I_2$  is the output, Eq. 29 can be reduced to

$$y = I_2 = p_{11}E_1 + p_{12}E_2 \quad (32)$$

$$z = I_1 = p_{21}E_1 + p_{22}E_2$$

A comparison of these equations with, for example, Eq. 276 of reference [8] shows that the p's are (with the proper change of signs and subscripts) identical to the ordinary "y" parameters used to describe a two-port communication network. In fact the p's are the short-circuit driving point and transfer admittances for the network. If  $I_1$  becomes the input and  $E_2$  the output, the resulting p's will be the "z" parameters, or open circuit impedances. Other choices of input and output variables will result in the "g" and "h" parameters.

The transfer matrix P is essentially only a generalization of these circuit parameters for multi-port networks. (The "ABCD" or general circuit parameters which express  $E_2$  and  $I_2$  in terms of  $E_1$  and  $I_1$ , while they are a perfectly accurate mathematical way of describing a system, cannot be considered a special case of the P matrix because the corresponding equations imply that the two conjugate variables at one terminal are being considered simultaneously independent.)

This example also serves to illustrate the fact that a general n-port bilateral network can be represented by  $2^n$  different transfer matrices,

depending upon the choice of dependent and independent variables. The P matrix therefore is determined not only by the physical system, but also by the way in which it is to be used.

### e. A Matrix Definition for Output Interaction

Since the outputs  $y_i$  have been considered dependent variables, we cannot discuss external forcing of these variables without destroying the validity of the essentially unilateral transfer matrix. However, if  $y_i$  is the position of output shaft i, if we apply enough torque to shaft i, we can produce whatever disturbance is desired in  $y_i$ . The effect of a disturbance at output shaft i is really the effect of the secondary input conjugate to  $y_i$ . If a signal  $w_i$  does not affect  $y_j$ , and a signal  $w_j$  does not affect  $y_i$ , then the outputs  $y_i$  and  $y_j$  are non-interacting. (In the position example, if  $y_i$  is obliged to take up a position externally, the torque  $w_i$  necessary to accomplish this does not disturb  $y_j$  and vice versa.)

The P-canonical structure of reference [2] is defined as a system where each output depends upon all inputs, but is not affected by disturbance to any other output. This condition is satisfied if matrix B of the preceding section is diagonal. The matrix A will be identical to the matrix P of reference [2].

In the V-canonical structure, every output interacts with every other output. In the matrix representation, B will have no non-zero terms. This condition means that a disturbance applied to  $w_i$ , the secondary input conjugate to  $y_i$ , will alter the values of all other  $y$ 's, as well as that of  $y_i$ .

The H-canonical structure is defined in reference [4] for a system having more outputs than inputs. If  $m - n = i$ , it is assumed that  $y_1 \dots y_i$  are not affected by output disturbances, but that  $y_{i+1} \dots y_m$  are affected by a disturbance in any output. This condition results in a B matrix of the following form:

$$B = \begin{bmatrix} P_1(n+1)^0 & \cdots & 0 \\ P_2(n+1) & P_2(n+2)^0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ P_1(n+1) & \cdots & P_1(n+1)^0 & \cdots & 0 \\ P_{(l+1)(n+1)} & \cdots & P_{(l+1)(n+l+1)} & P_{(l+1)(n+m)} & & & & & & & & & & & & & \\ \vdots & \vdots & \ddots & \\ P_m & \cdots & P_n(n+m) \end{bmatrix} \quad (33)$$

Note that while the V-canonical structure is limited to systems having the same number of outputs as inputs, there is no reason to require the presence of zeroes in matrix B, regardless of the relation between m and n. The P matrix always relates  $n + m$  dependent variables to  $n + m$  independent variables, and it is a matter of physical and mathematical indifference whether n equals m. Completely interrelated systems (those where a disturbance at any output terminal is reflected in all outputs) can certainly exist, even though the number of outputs is greater than the number of inputs.

In the case of the P-canonical structure, the matrix B gives exactly the same information as the assumption about structure, viz., the outputs are completely non-interrelated in the sense considered here.

In the V-canonical structure the block diagram does not give quantitative information about the effect at  $y_i$  of a disturbance  $y_j$ . If, however, in the P matrix representation we let all independent variables (x's and w's) be zero except  $w_j$ , then

$$y_i = P_i(n+j)w_j \quad (34)$$

But also

$$y_j = P_j(n+j)w_j$$

Therefore, the effect on  $y_i$  of a disturbance at the  $j$ th terminal producing an output  $y_j$  is  $y_i = (p_i(n+j)/p_j(n+j)) y_j$ .

The fact that both  $y_i$  and  $y_j$  are regarded as the effects of a single cause  $w_j$  does not alter the conclusion that an output  $y_i$  corresponds to a disturbance  $y_j$ .

The P matrix can be completely determined externally for an unknown system only if all independent variables can be successively manipulated and all dependent variables simultaneously observed. However, the degree to which P must be known depends upon the application; if some of the secondary outputs cannot be observed, the corresponding rows may be omitted from the P matrix. Similarly, if some of the secondary inputs are always zero in a given application, the corresponding columns may be omitted.

#### f. Bilateral versus Unilateral Matrices

In Section III c, it was stated that for an electric network, the P matrix can be derived from Kron's Z' matrix, and in reference [7] it is shown that for a network of the type considered, Eq. 2 can always be solved for the unknown quantities, whichever they may be. This is equivalent to saying that whatever variables are taken to be inputs and outputs (subject to the constraint that the current and voltage at one node pair are not both inputs or outputs) the corresponding P matrix can be found.

But the P matrix represents  $n + m$  equations in  $2(n + m)$  variables; if the specification of any  $n + m$  uniquely determines the other  $n + m$  (and the above statement implies that it does for electric networks or their analogs), then the equations represented by P must hold regardless of which set of variables is considered to be unknown. In other words, to go from the P matrix representing one set of inputs to the P matrix of any other set, it is only

necessary to solve the equations for the new unknowns. Thus, although Eq. 29 was written under the assumption that the terms of  $V$  were results (that is, determined by the independent variables  $U$  and the system equations), it continues to be valid even if some of the  $u$ 's become dependent, and some of the  $v$ 's independent. For example,  $V = P^{-1} U$ .

In a sense, then, unless the system represented by  $P$  cannot be simulated by an electric network, the  $P$  matrix description is bilateral — its equations are valid regardless of the choice of input and output variables. The  $A$  matrix, being only a part of the  $P$  matrix, is not bilateral.  $Y = AX$  implies that  $W = 0$ , while  $X = A'Y$  implies that  $Z = 0$ . Hence  $A'$  is not generally equal to  $A^{-1}$ . It is, at least in many cases, this reason which causes single-variable transfer functions to be limited to signals passing in one direction.

#### g. State Vector Method; Controllability and Observability

The  $P$  matrix, while providing information regarding output interaction, suffers from the same disadvantages as the  $A$  matrix as far as initial conditions are concerned. The state vector method, on the other hand, is particularly suited to problems involving initial conditions such as those arising in optimal control theory. For time-varying systems where Fourier transform methods fail and non-linear systems where the superposition principle is invalidated, the designer has recourse only to the state vectors or differential equation method of representation.

By definition, when the system is specified by equations of the form

$$\dot{x} = Ax + Bu \quad (35)$$

$$y = Cx + Du$$

the system is assumed to be completely defined by the "n" state variables.

This implies that the designer has access to all points within the system or

a priori knowledge of the order of the system. If every state of the system is controllable then the state vector can be transferred from a given state to any other desired state. It must be remembered that any non-singular linear transformation  $Z = PX$  yields a different but equivalent set of state variables  $z_i$ .

For the time-invariant situation such a transformation does not affect stability or controllability of the system. While the state vector representation is a mathematical model of the system which describes it completely it is not unique and the state variables and consequently the equations themselves are chosen to simplify manipulations.

Like the A matrix, the state vector method is unidirectional in character

$$X(t) = \Phi X_0 + \int_{t_0}^t \Phi(t, \tau) B(\tau) U(\tau) d\tau. \quad (36)$$

$X(t)$  may be determined from the initial condition  $X_0$  and the input  $U(t)$  to the system. Every state variable  $x_i$  is a dependent variable and may be affected only by changing the input  $U(t)$ . Hence, output interaction interpreted as the effect of varying one output on any other output cannot be described using this representation. On the other hand, the homogeneous solution  $\Phi X_0$  gives complete information regarding the effect of an initial condition of state variable  $x_i$  on the response of the system in the absence of an input  $U(t)$ .

A single-variable system which can be represented by an nth-order differential equation may also be expressed in state vector notation in the form shown on the following page, where  $x_1$  is the output of the system and  $x_2 \dots x_n$  are its  $(n - 1)$  derivatives. In the multivariable case, the state vector  $X(t)$  consists of  $n$  components of which only  $q$  are the actual outputs of the system.

while  $(n - q)$  are secondary state variables, some of which are linear combinations of the derivatives of those outputs. The remaining ones represent unobservable state variables which do not affect the terminal behavior of the system.

Under zero initial conditions, the state vector representation may be converted into an  $(n \times r)$  matrix of transfer functions by the simple process of taking transforms. Of this, a  $(q \times r)$  submatrix is the same as the  $A$  matrix of the system.

When the latter is specified and the system has no unobservable state variables, the remaining  $(n - q)$  rows of the  $(n \times r)$  matrix are determined by the arbitrary choice of the secondary state variables.

$$\dot{x} = Ax + Bu$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad A = \begin{bmatrix} 0 & 1 & & & & \\ 0 & 0 & 1 & & & \\ 0 & & 0 & \ddots & & \\ & & & 1 & & \\ -a_0 & -a_1 & \cdots & -a_{n-1} & & \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

When the system has unobservable state variables the transfer matrix  $A$  contains no information about them. The actual order of the system cannot be predicted using any transfer function representation. In other words, when the transfer matrix is used it is tacitly assumed that unobservable and uncontrollable state variables are unimportant. Problems in which controllability and observability do play an important part must be handled by the state vector notation.

#### IV. Conclusion

The aim of the foregoing discussion has been a clarification of the limitations and advantages of various methods of describing a multivariable system. The results can be summarized in the following statements.

(1) The conventional transfer matrix, or A matrix, describes the terminal behavior of a linear, time-invariant system; this description is valid for a given set of input and output variables. The A matrix provides no information about the response of the system when an external signal is applied at any output. It therefore does not define output interaction. Since the transfer functions are obtained by dropping terms in the transformed differential equation containing the effect of initial conditions, properties like controllability and observability must be considered irrelevant for systems represented only by a transfer matrix.

The A matrix is thus best suited to problems in which the steady state behavior or transient response with zero initial conditions is of primary importance, and where only terminal variables are of interest.

(2) The division of a system represented by a transfer matrix into single-variable subsystems described by transfer functions does not provide any additional information about terminal behavior. In particular, output interaction cannot be defined by means of such subdivision. Exactly as in the case of single-variable systems, all configurations of subsystems are equivalent mathematical models of the physical system; the choice of a particular configuration depends on convenience of manipulation.

A similar method of defining a mathematical structure is the selection of a particular set of state variables; an infinite number of state vector representations are possible for any linear time-invariant system and the choice of any one of these is again a matter of convenience.

(3) Problems involving signal flow in two directions at a given terminal can be studied by means of the expanded transfer matrix, or P matrix. This matrix, while it shares the inability of the A matrix to take into account

non-zero initial conditions, can be useful in situations where output interactions or non-isolated interconnections of systems are important.

(4) Of the methods considered in this report, the state vector representation is the only way of describing time-varying or non-linear multivariable systems, since the frequency domain methods are no longer valid. This approach provides complete information regarding the response to an input vector  $U$  in the presence of initial conditions. It is most suitable for problems in which system behavior is to be controlled directly in the time domain, and in the area of optimal control. The use of digital computers for obtaining approximate solutions makes this method especially attractive. However, this method cannot be used for systems with distributed parameters because the number of state variables becomes infinite.

(5) The number of initial conditions that have to be specified to obtain a unique solution depends on the order of the system, and is equal to the number of state variables. When the system is described only by a transfer matrix, however, the order of the system is not obvious. Hence, in situations where every state of the system must be independently controlled, only the differential equation representation should be used. It is only in such situations that controllability and observability are important.

References

1. H. Freeman, "A Synthesis Method for Multipole Control Systems," AIEE Transactions, Vol. 76, Part II, pp. 28-31 (1957).
2. M. D. Mesarovic, The Control of Multivariable Systems, John Wiley and Sons, New York, N. Y., 1960.
3. A. S. Boksenbom and R. Hood, "General Algebraic Methods Applied to Control Analysis of Complex Engine Types," Technical Report 980, NACA, Washington, D. C. (April, 1949).
4. R. E. Kalman, "On the General Theory of Control Systems," Proc. First International Congress of Automatic Control, Moscow, USSR (1960).
5. E. G. Gilbert, "Controllability and Observability in Multivariable Control Systems," to be published.
6. R. E. Kalman, Y. C. Ho, and K. S. Narendra, "Controllability in Linear Dynamical Systems," Contributions to Differential Equations (1961).
7. P. Le Corbeiller, Matrix Analysis of Electrical Networks, Harvard University Press, Cambridge, Mass., 1950.
8. E. A. Guillemin, Communication Networks, Vol. II, John Wiley and Sons, New York, N. Y., 1935.

110019990000200 0400000000000000

**Washington State College**  
Pullman, Washington  
**West Lecture, June 1940**  
  
**Washington Economic Commission**  
Seattle, Washington  
**Washington State University at Pullman**  
111 Palouse Street  
Pullman, Washington  
  
**Mr. L. B. Shambaugh**  
Seattle, Washington  
**Washington State Laboratories**  
Olympia City, Washington  
  
**Professor E. W. Byrnes**  
P. O. Box 200  
Bellevue, New Jersey